EXHIBIT
D

Making AI systems you can rely on

Anthropic is an AI safety and research company. We build reliable, interpretable, and steerable AI systems.

Join Us

Our Purpose   **We believe AI will have a vast**

impact on the world. Anthropic is dedicated to building systems that people can rely on and generating research about the opportunities and risks of AI.

We Build Safer Systems

We aim to build frontier AI systems that are reliable, interpretable, and steerable. We conduct frontier research, develop and apply a variety of safety techniques, and deploy the resulting systems via a set of partnerships and products.

Safety Is a Science

We treat AI safety as a systematic science, conducting research, applying it to our products, feeding those insights back into our research, and regularly sharing what we learn with the world along the way.

Interdisciplinary

Anthropic is a collaborative team of researchers, engineers, policy experts, business leaders and operators, who bring our experience from many different domains to our work.

AI Companies are One Piece of a Big Puzzle

AI has the potential to fundamentally change how the world works. We view ourselves as just one piece of this evolving puzzle. We collaborate with civil society, government, academia,

The Team

We're a team of researchers, engineers,
policy experts and operational leaders,
with experience spanning a variety of
disciplines, all working together to build
reliable and understandable AI systems.

Research

We conduct frontier AI research across a variety of modalities, and explore novel and emerging safety research areas from interpretability to RL from human feedback to policy and societal impacts analysis.

Policy

We think about the impacts of our work and strive to communicate what we're seeing at the frontier to policymakers and civil society in the US and abroad to help promote safe and reliable AI.

Product

We translate our research into tangible, practical tools like Claude that benefit businesses, nonprofits and civil society groups and their clients and people around the globe.

Operations

Our people, finance, legal, and recruiting teams are the human engines that make Anthropic go. We've had previous careers at NASA, startups, and the armed forces and our diverse experiences help make Anthropic a great place to work (and we love plants!).

Our Values

01
Here for the mission

Anthropic exists for our mission: to ensure transformative AI helps people and society flourish. Progress this decade may be rapid, and we expect increasingly capable systems to pose novel challenges. We pursue our mission by building frontier systems, studying their behaviors, working to responsibly deploy them, and regularly sharing our safety insights. We collaborate with other projects and stakeholders seeking a similar outcome.

04
Do the simple thing that works

We celebrate trying the simple thing before the clever, novel thing. We embrace pragmatism - sensible, practical approaches that acknowledge tradeoffs. We love empiricism - finding out what

02
Unusually high trust

Our company is an unusually high trust environment: we assume good faith, disagree kindly, and prioritize honesty. We expect emotional maturity and intellectual openness. At its best, our trust enables us to make better decisions as an organization than any one of us could as individuals.

03
One big team

Collaboration is central to our work, culture, and value proposition. While we have many teams at Anthropic, we feel the broader sense in which we are all on the same team working together towards the mission. Leadership sets the strategy, with broad input from everyone, and trusts each piece of the organization to pursue these goals in the best way they see. Individuals

actually works by trying it - and apply this to our research, our engineering and our collaboration. We aim to be open about what we understand and what we don't.

commonly contribute to work across many different areas.

Company News

Thoughts on the US Executive Order, G7 Code of Conduct, and Bletchley Park Summit

Nov 5, 2023 · 4 min read

Dario Amodei's prepared remarks from the AI Safety Summit on Anthropic's Responsible Scaling Policy

Nov 1, 2023 · 5 min read

Want to help us build the future of

AI?

Join Us

Product

Research

Index

Company

News

Careers

Press Inquiries

Support

Twitter

LinkedIn

Terms of Service

Privacy Policy

Acceptable Use
Policy

Responsible
Disclosure Policy

Compliance

© 2024 Anthropic PBC